

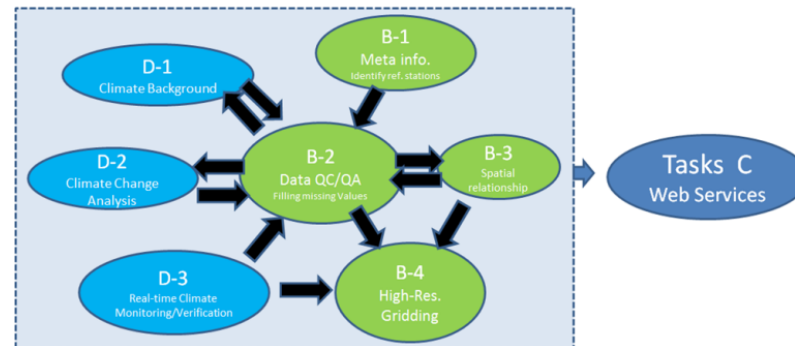
# 台灣自動氣象站氣溫資料補遺方法探討及網格化分析

陳雲蘭<sup>1</sup> 陳品妤<sup>1</sup> 詹智雄<sup>1</sup> 沈里音<sup>1</sup> 馮智勇<sup>2</sup> 劉家豪<sup>2</sup> 林佑蓉<sup>2</sup>

<sup>1</sup>中央氣象局

<sup>2</sup>多采科技有限公司

“臺灣長期氣候資料整集、處理及分析”  
今(2014)年度工作規劃



B: 建構日後可提供完善資料服務的能力

D: 建構日後可提供完善氣候分析資訊服務的能力

Start from Temperature....

# 氣候變遷應用服務能力發展計畫 --氣候資料整集分析系統發展

## 強化檢核測站資料、持續校驗網格化技術

### 測站詮釋資訊

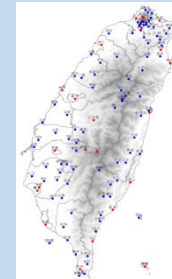
B-1

Meta info.  
Identify ref. stations

年度目標：完成站況分析報告--調查本局的**長期氣象觀測資料**

- 先以溫度為指標要素，依不同應用目的分類，根據資料具一定程度的完整性(達85%)，整理出應用目標參考測站清單，以作為後續各類資料處理及分析的依據。
- 站況調查：除了基本資訊，調查重點在於是否有遷站，新舊接站，觀測(儀器)改變，重要記事。以提供後續資料分析之必要資訊。

### 網格化參考測站(111站)



### 邁向均一化

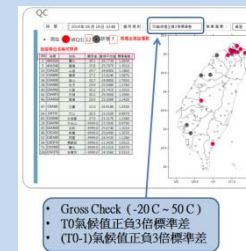
B-2

Data QC/QA  
Filling missing Values

年度目標：完成一套**經QC及補遺**等整理的氣溫觀測資料

- 以小時資料為單位進行缺失記錄調查。
- 首年先以簡易但可信之方法處理QC及缺補，同時調查及測試WMO等組織介紹之均一化套件，待明年再實現均一化的檢測及處理。

### 設計異常資料警示輔助分析圖表



### 邁向網格化

D-3

Real-time Climate  
Monitoring/Verification

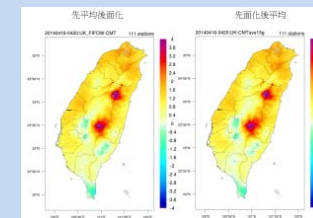
B-4

High-Res.  
Gridding

年度目標：完成一套**氣溫高解析網格資料**，並說明方法及不確定性。

- 測試不同克利金參數方式的差異，同時以另一種網格化(美國使用的BCDG)方法提供比對。
- 使用至少2套不同方法逐月產出近期高解析網格化氣溫資料，以了解不同網格化方法的差異程度，並幫助提供監測資訊。
- 決定對外服務資料的網格化採行方式，以經過QC及補遺的近15年資料，產製網格資料，並包含說明文件。

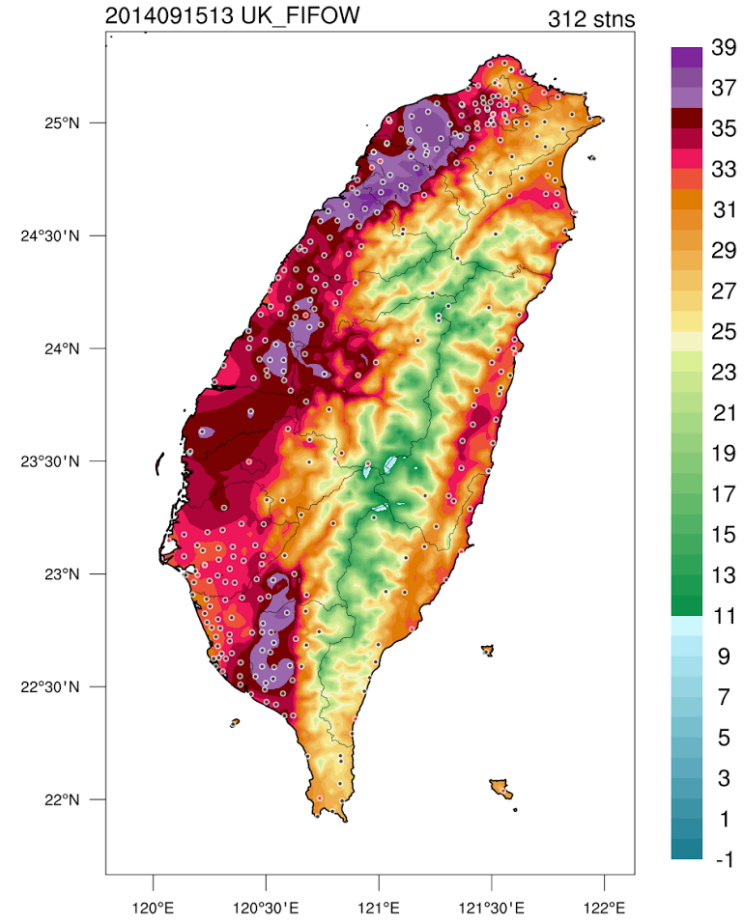
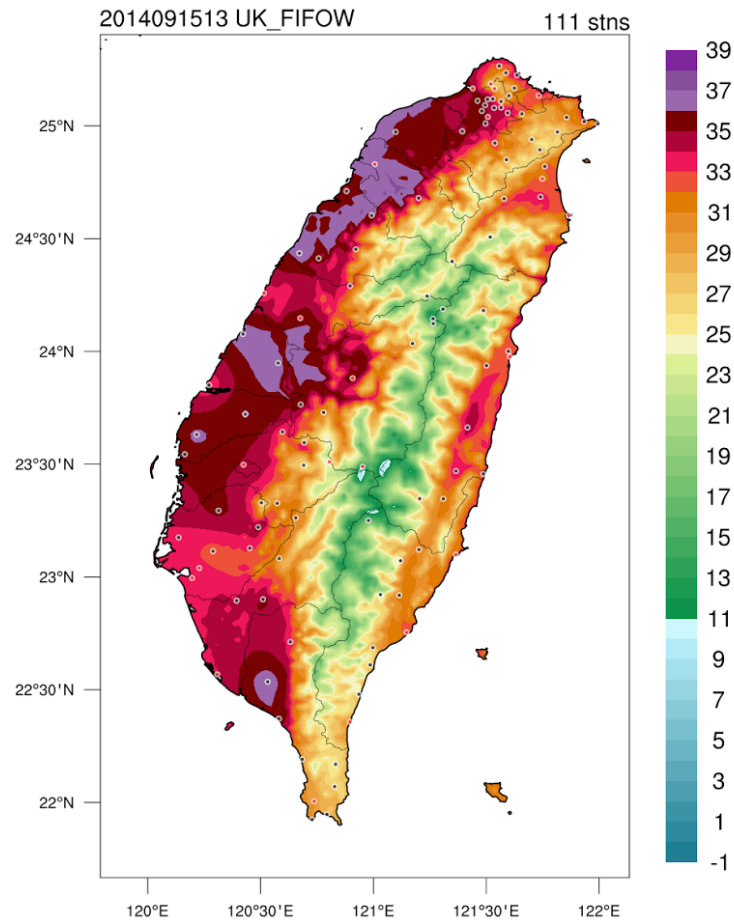
### 資料處理程序的差異比較



# 網格化推估值受參考測站資料品質及測站數量多寡影響

選用111個參考站

使用現有312個參考站



2014年9月15日下午桃竹苗極端高溫事件氣溫分布圖

# 台灣自動氣象站氣溫資料補遺方法探討及網格化分析

## 報告內容

- 氣溫資料網格化參考測站的選取方式
- 克利金空間統計法在氣溫資料網格化的可信度分析
- 統計資料採樣均一化處理
  - 撤遷站資料處理方法
  - 觀測缺失資料補遺方法
  - 測站氣溫資料補遺結果比對

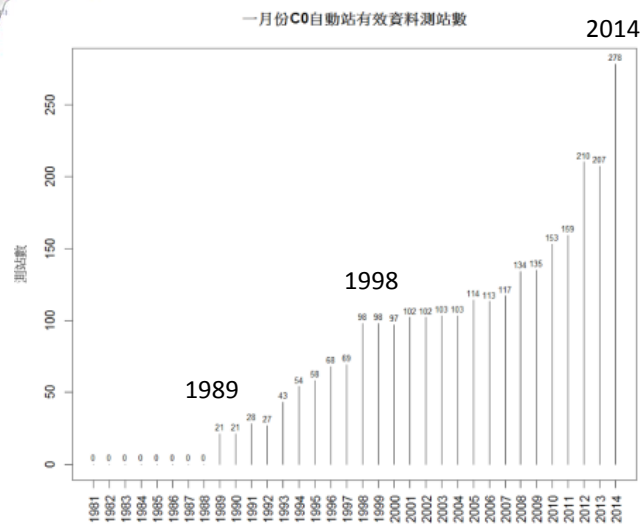
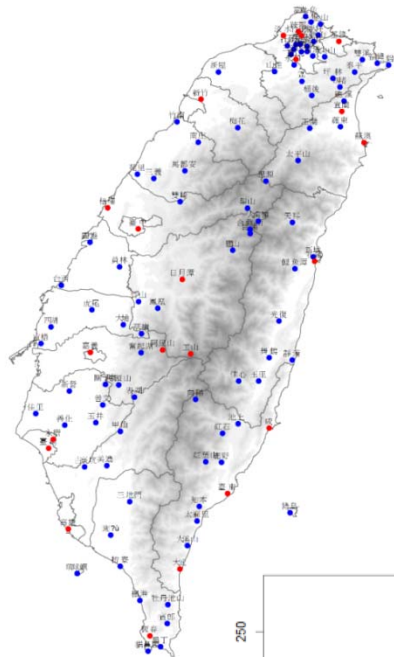
本報告著重在資料缺失補遺的處理，  
使用方法依賴同樣用於網格化的克利金空間統計法。

克利金方法介紹及應用, 請詳見：

- 李天浩，2009：應用克利金法建立高解析度網格點氣象數據之研究。交通部中央氣象局委託研究計畫成果報告。
- 馮智勇、劉家豪、陳雲蘭，2012：客觀分析法地面溫度案例分析。天氣分析與預報研討會論文彙編，中央氣象局。
- Yun-Lan Chen et al. 2014, Understanding Uncertainty in the Taiwan High-Resolution Gridded Dataset of Monthly Surface Temperature. 2014 TCCIP International Workshop on Climate Change

## 網格化參考站(具長期資料)的選擇

### 高解析網格化分析參考站(22+89 = 111)

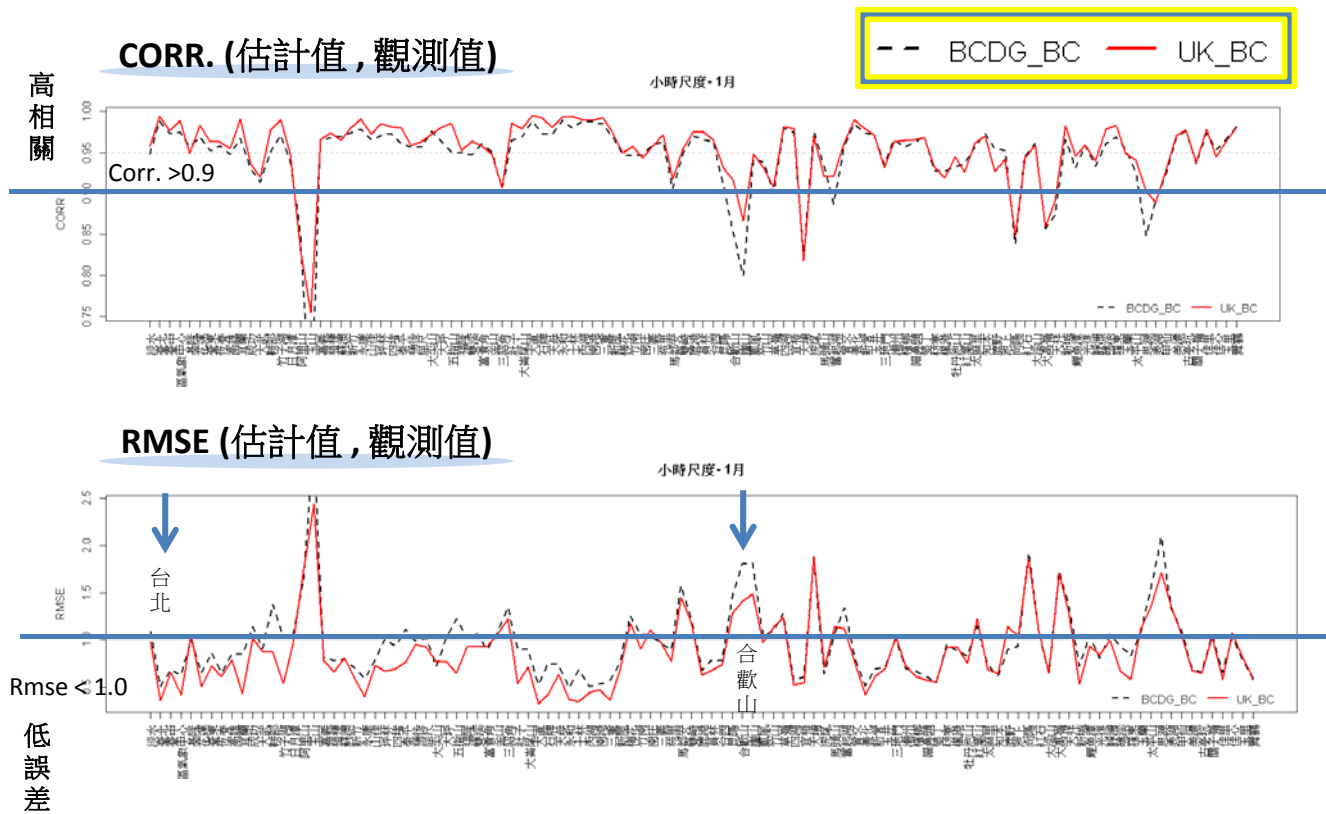


- 初期先只取人工站(46xxx)及自動氣象站(COxxx)。
- 先以1月份為根據，統計1998~2012期間的時資料量，列出有效值達85%者。
- 雖然各月份資料量有差異，初期將在各月的分析暫時先一致以此111站為目標參考測站。
- 後續分析過程仍將持續檢討目標參考站的取捨。

逐月正常資料站數	(正確資料者/全部資料)%			
	n=95%	n=90%	n=85%	n=80%
Jan	91	106	111	111
Feb	87	106	111	111
mar	84	103	110	111
apr	76	101	108	111
may	78	101	109	110
jun	86	99	109	111
jul	77	97	108	110
aug	71	91	107	110
sep	80	95	106	111
oct	72	98	107	110
nov	85	102	107	111
dec	95	105	107	110

# 克利金空間統計法推估氣溫可信度分析

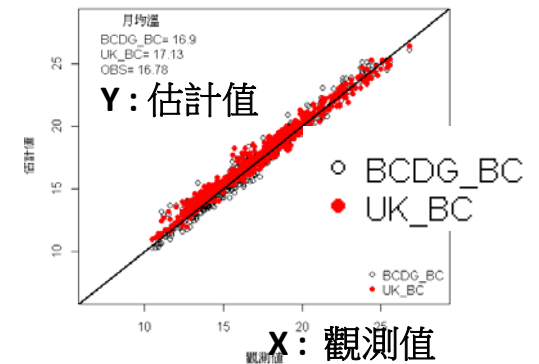
- 了解採用通用克利金(UK)法對各測站氣溫推估的不確定程度。
  - 以逐一遮蔽測站法，分析各站估計值與實測值的相關程度及誤差。
  - 同時比對另一內插工具BCDG法(以Cressman為基礎)的估計成效。



## 時資料估計值散布圖

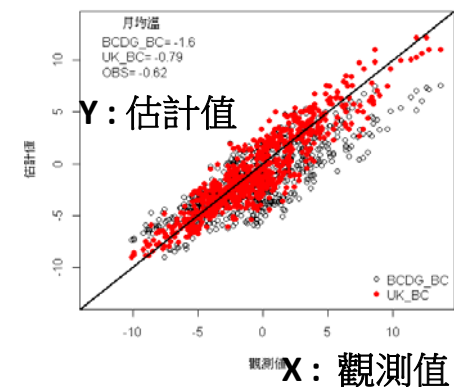
### 平地站：以台北為例

臺北-2014 1月



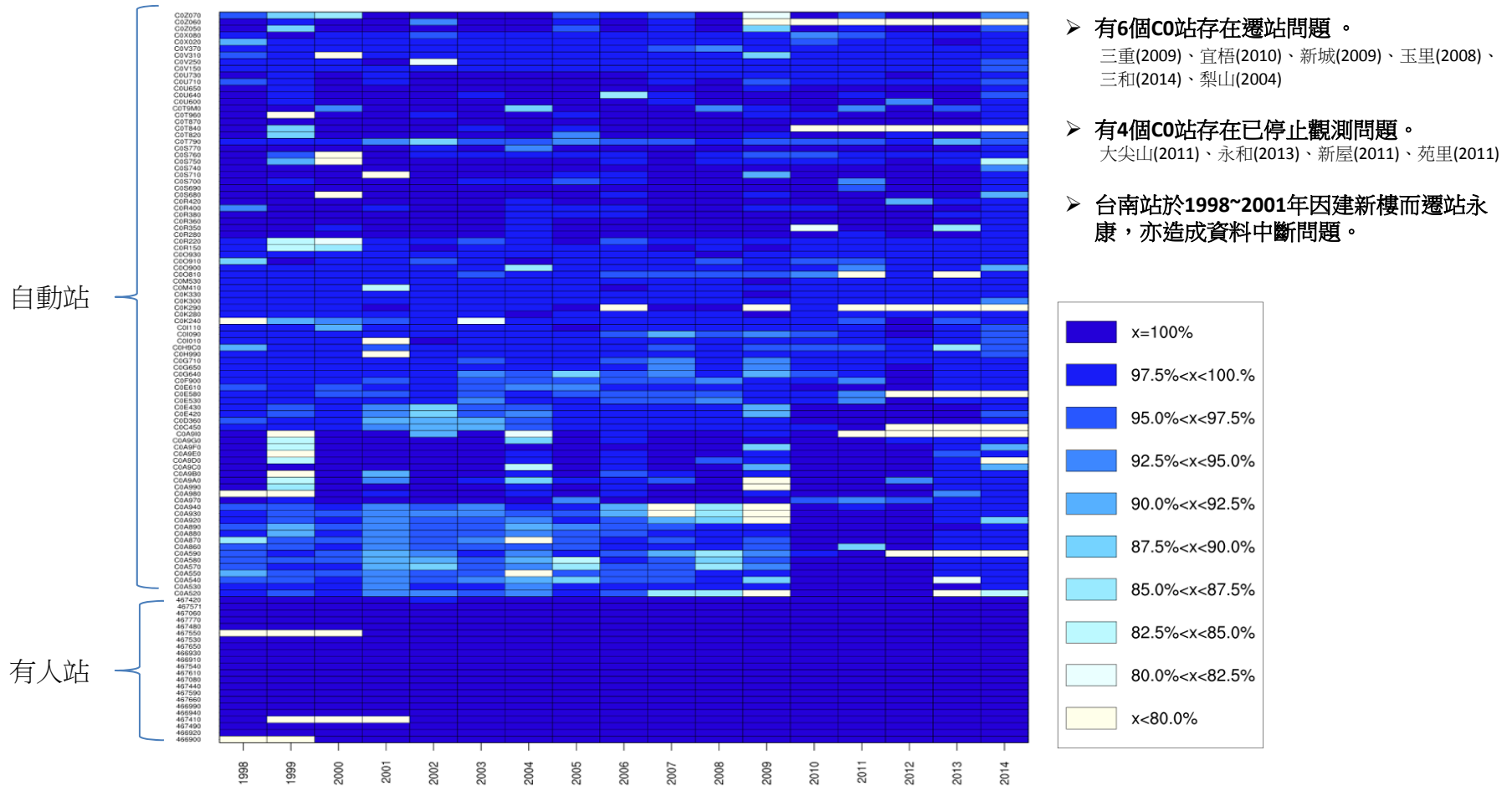
### 山區站：以合歡山為例

合歡山-2014 1月



# 統計資料採樣均一化處理

為避免資料樣本數的不一致，我們建議在網格化程序之前需對資料進行完整補遺。工作包含：(1) 中斷資料的介接 (2) 缺漏觀測的補遺。



111個網格化參考測站觀測資料齊整率分布圖 (1998~2014,Jan.)

## (1) 介接撤遷站資料

---

依二類情形分別處理撤遷測站的資料中斷問題：

### 1. 有鄰近替代站：**UK**

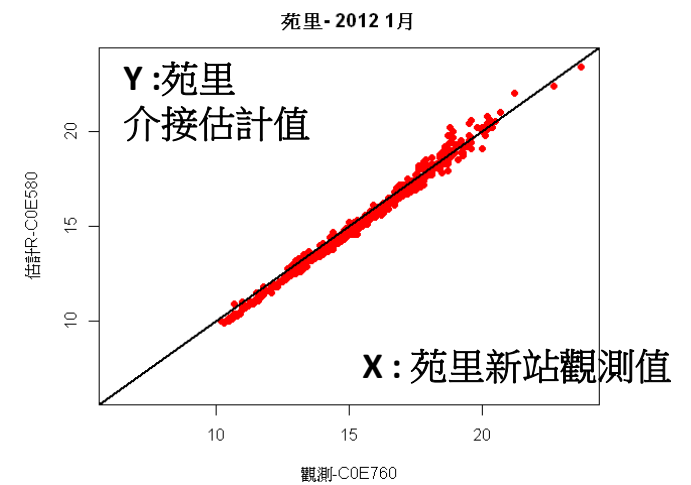
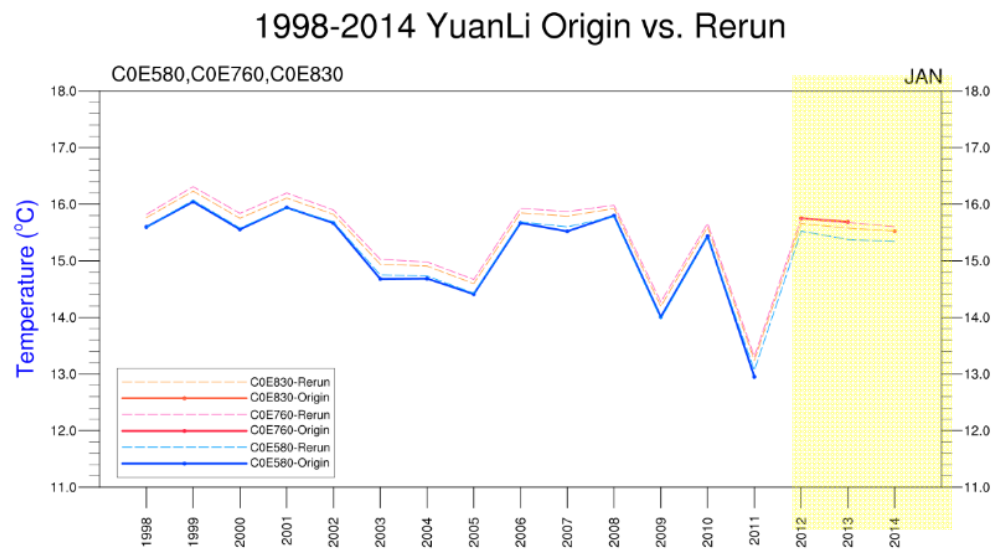
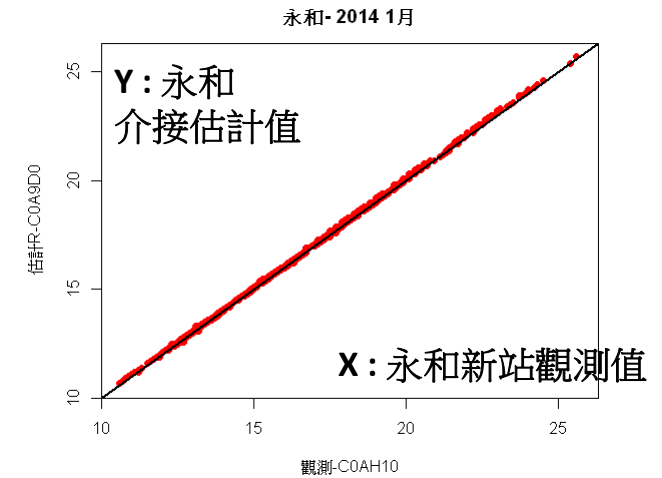
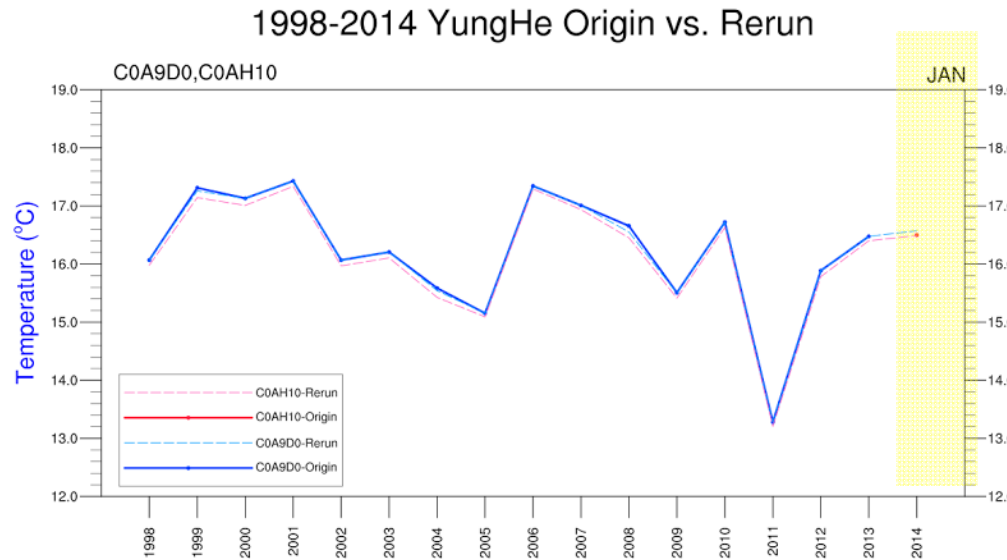
- 利用克利金空間統計法來求算舊站地點估計值，此時鄰近替代站通常將反應最大權重。
- 目前所需處理的11個撤遷測站中有7個站(三重、苑里、宜梧、新城、玉里、三和、永和)用此方法完成新舊站資料介接。

### 2. 沒有鄰近替代站：**UK\_TakeOneOut + BiasCorrection**

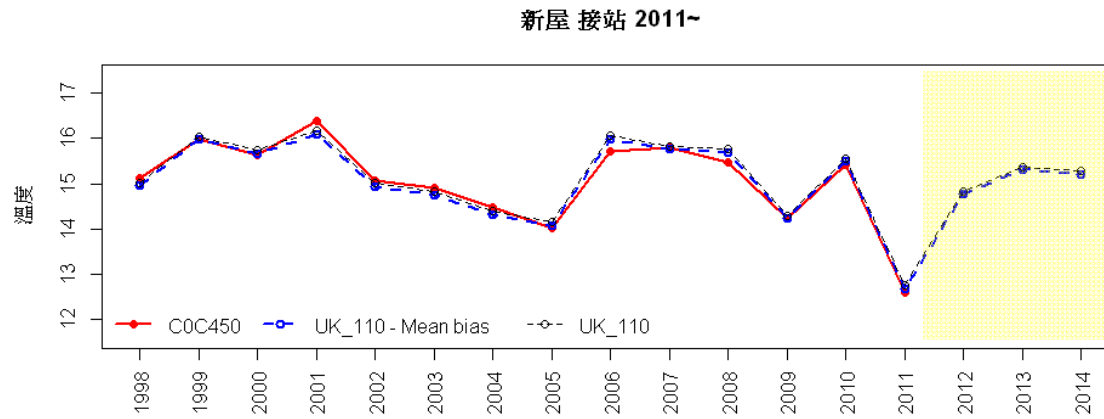
- 仍使用空間統計法來求算因撤站或資料中斷期間的估計值，但因為沒有位置接近的鄰近參考站，若此待處理測站具有非其他測站可解釋的局地氣候特徵(例如與地表特徵有關的都市效應)，將影響推估正理性。因此在過程中另外輔以調整系統性偏差值，乃利用過去使用所有參考站對此缺遺站的估計平均偏誤分析(遮蔽測站估計法)來決定調整值。
- 目前所需處理的撤遷測站中，台南、梨山、大尖山、及新屋4個站用此方法完成資料中斷的問題處理。



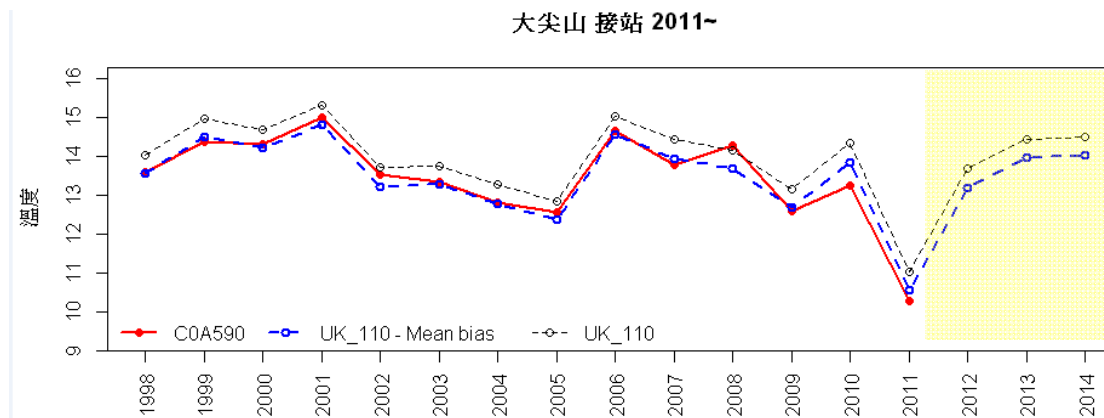
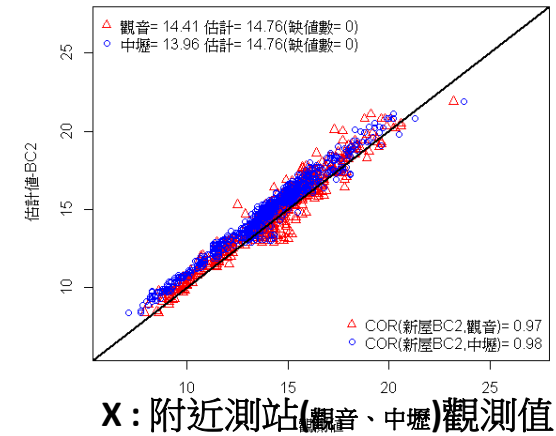
有鄰近替代站的資料介接：以永和(上圖)、苑里(下圖)為例



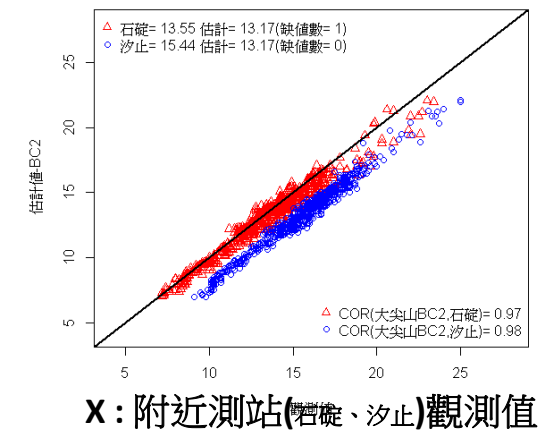
# 沒有鄰近替代站的資料交接：以新屋(上圖)、大尖山(下圖)為例



**Y: 新屋**  
交接估計值 新屋-2012 1月



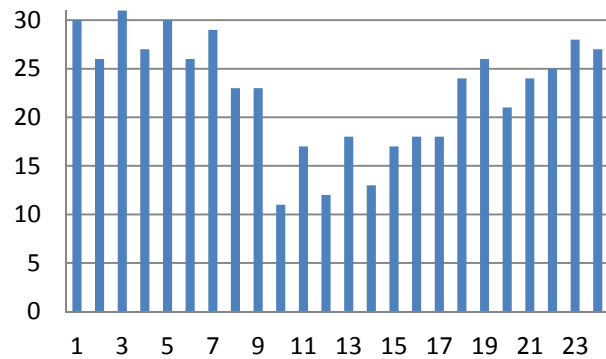
**Y: 大尖山**  
交接估計值 大尖山-2012 1月



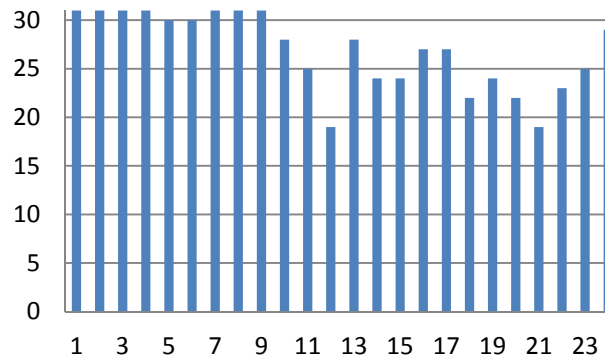
對於統計採樣均一問題，除了介接中斷資料(僅有少數站)，還要補漏(存在於多數站)！

缺遺值在24個小時分布不均，將影響對日、月統計量的合理代表性

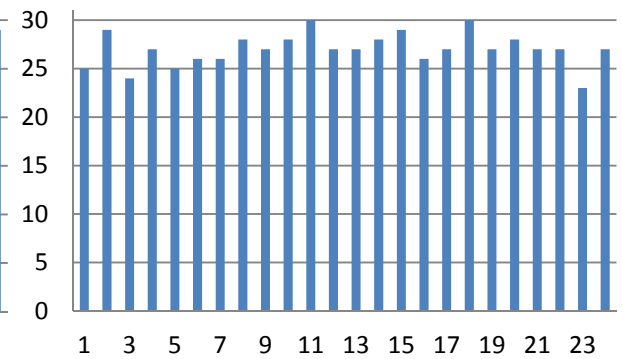
山佳2014七月



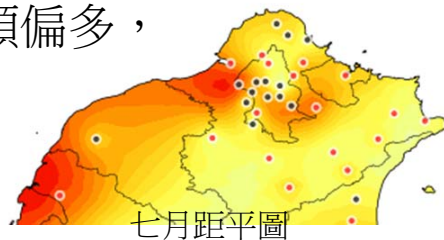
天母2014七月



南庄2014七月



山佳站白天缺遺值明顯偏多，  
月平均值可能低報。



天母站上半夜資料較齊全，  
白天缺遺值多，  
如此統計出的月平均值亦可能低報。

## (2) 缺漏觀測的補遺

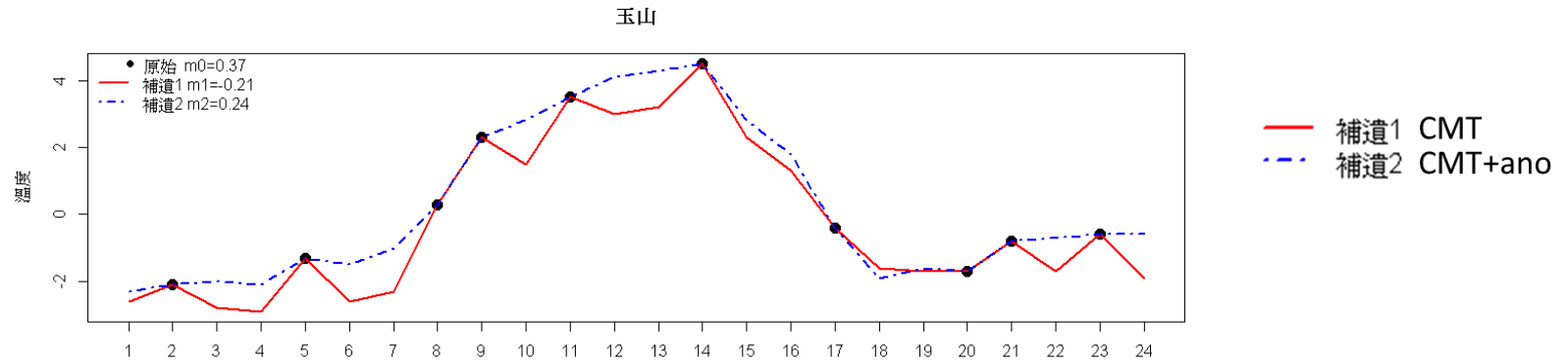
---

對於不齊備資料的處理，過去研究大多以前後資料求取平均處理，亦有以氣候值填補的做法。本計畫認為對於只有1~2筆的非連續性缺值，以前後筆觀測數值插補的方式確實可行，但對於更長時間的連續性缺值狀況，因單一站點的可用資訊已缺失，則需利用其他的資訊來源，例如透過與鄰近站相關程度來插補，或是利用動力模式分析的大氣資料以統計降尺度方法進行地面氣象數據推估。

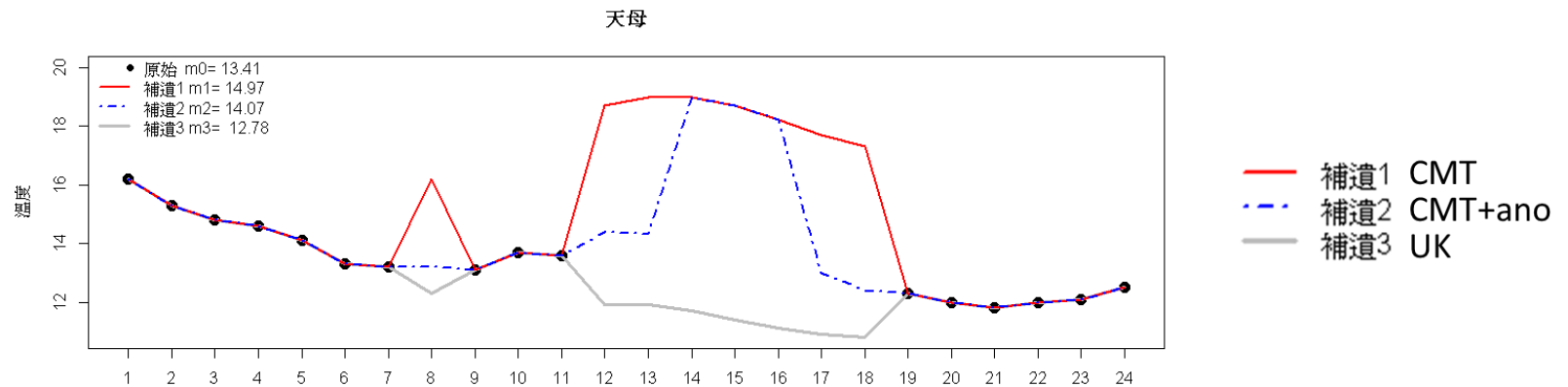
我們設計以下4種補遺方法進行比較：

- **Fill\_m1 (氣候值)**：直接以15年氣候平均值填補，以73個候作為時間樣本分類依據。
- **Fill\_m2 (氣候值 + 距平)**：仍利用氣候平均值，但加入對於當日天氣冷暖距平的考量。為改進Fill\_m1的做法，對於只有1~2筆的非連續性缺值，除了以氣候值填補並加入前後2筆觀測點的距平值，
- **Fill\_UK + Bias Correction**：乃為空間統計關係的插補做法。我們認為此法可以更好處理連續缺值的狀況。直接用Fill\_UK有可能不能反應出某些地方的局地特性，例如在冬季或夜晚時間，如果缺失值出現在市郊(例如淡水)，則以大多為市區的測站資料來推估此郊區站，雖然仍可有機會一定程度地掌握合理變化趨勢，但亦可能出現偏暖的高估值。面對這種偏估可能的問題，我們在方法設計中另外先以去一法來求出推估各站的平均偏估值，如此所得**空間統計關係的補遺值經過除去偏估有良好成效**，可大幅改進對長時間連續缺值的問題，彌補Fill\_m2所不能處理的問題。
- **Fill\_MOS**：另一種獲取局地特徵資訊的方式是利用NWP資料，利用統計降尺度技術與地面站資料透過迴歸模型建置，可以統計模型得出缺失值的推估值。

氣候值及距平資訊的簡易插補法(Fill\_m2)，可用於非連續長時間缺值的補遺。

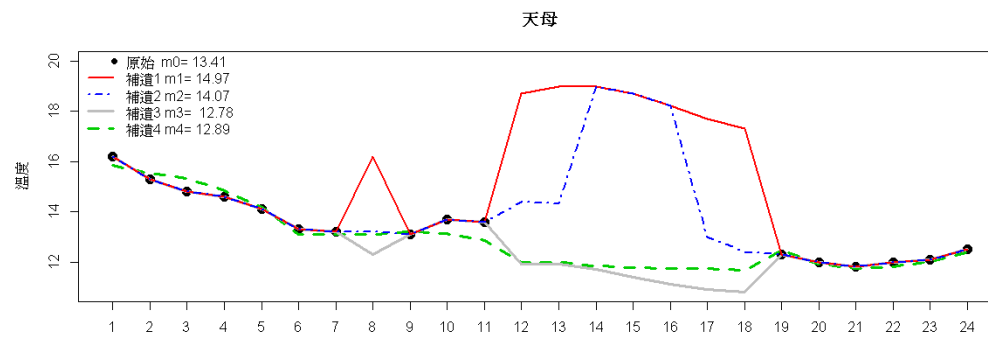


空間關係補遺法(UK)的優勢是可較好處理自動站許多連續性缺值的問題

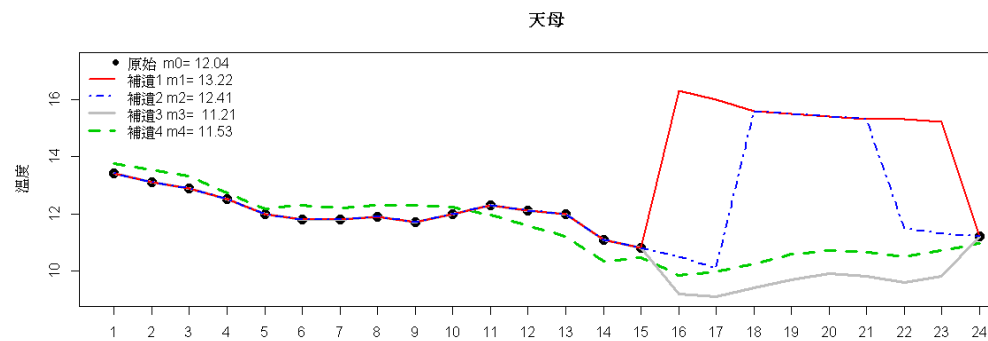


- 簡單利用多年氣候值加入缺值前後偏冷或偏暖的調整值，非常合適用於單一小時缺失值，也有機會對2筆少量缺值做合理推估。
- 至於更長的連續多筆缺值，因為已缺乏足夠鄰近的距平參考，僅能像Fill\_m1一樣直接以氣候值填補，若在此連續缺值時段剛好碰上比較波動的天氣，將因天氣偏離氣候值而使Fill\_m2填補值失宜。

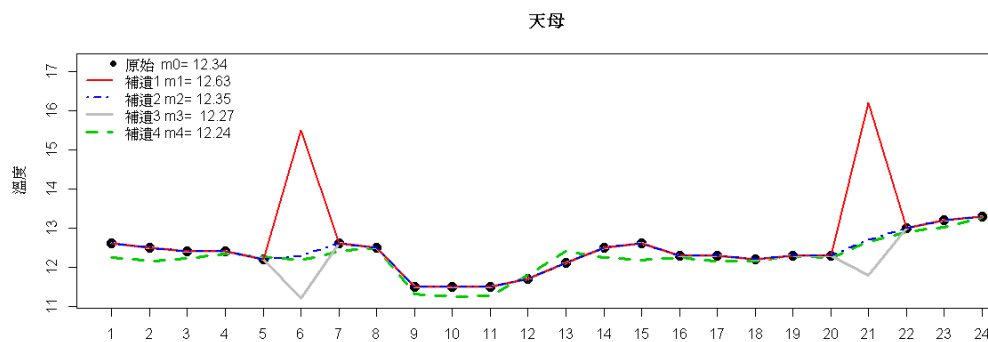
目前分析案例顯示：經過平均偏誤修正的空間關係補遺法有更佳表現



20040119



20040127



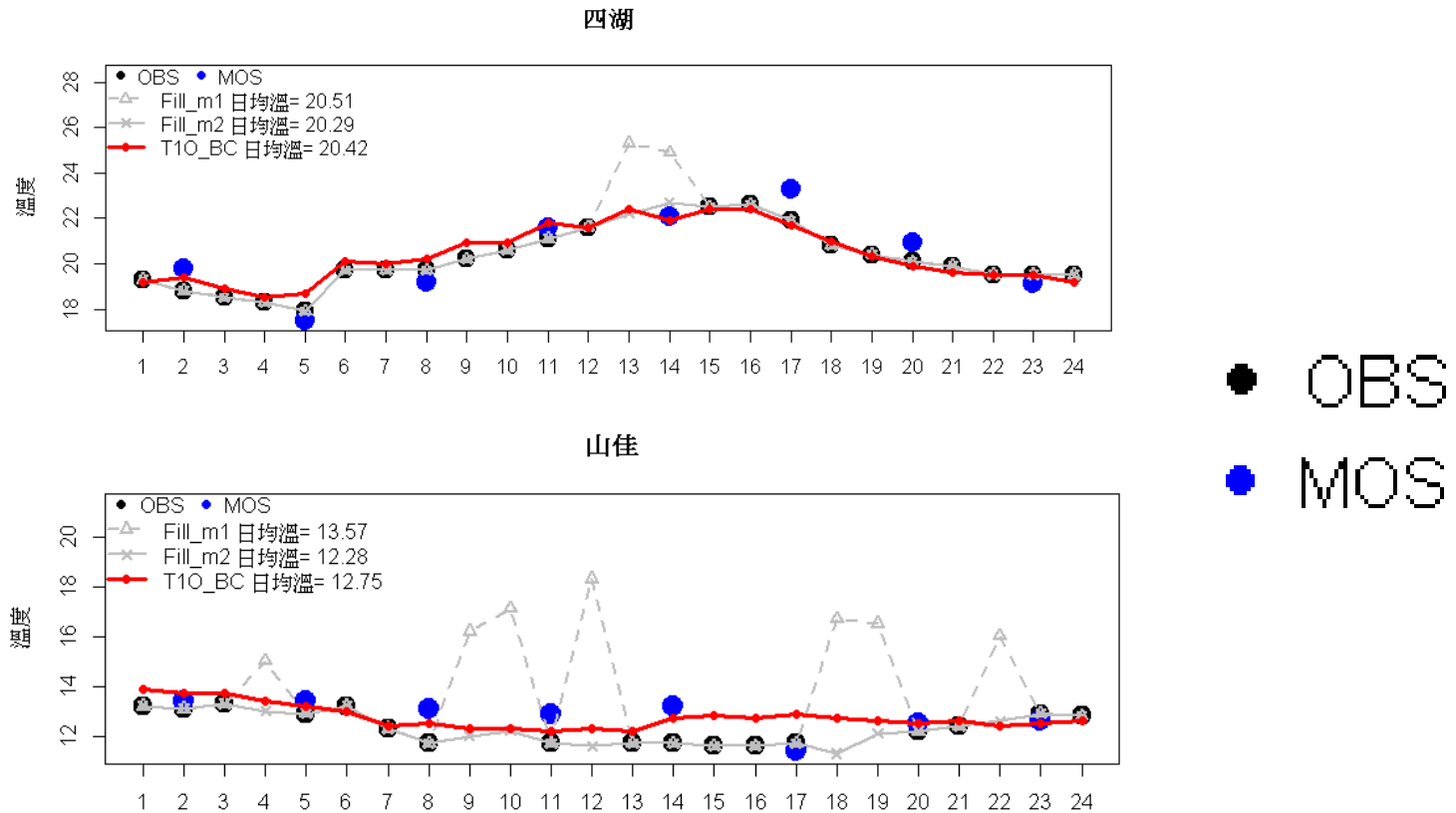
20040113

- 補遺1 CMT
- - - 補遺2 CMT+ano
- 補遺3 UK
- - - 補遺4 UK\_remove bias

## 4 種補遺方法結果比對

### 簡單補遺法(cmt+ano) VS 空間補遺法 (UK : T10\_BC)

#### 範例 (1)

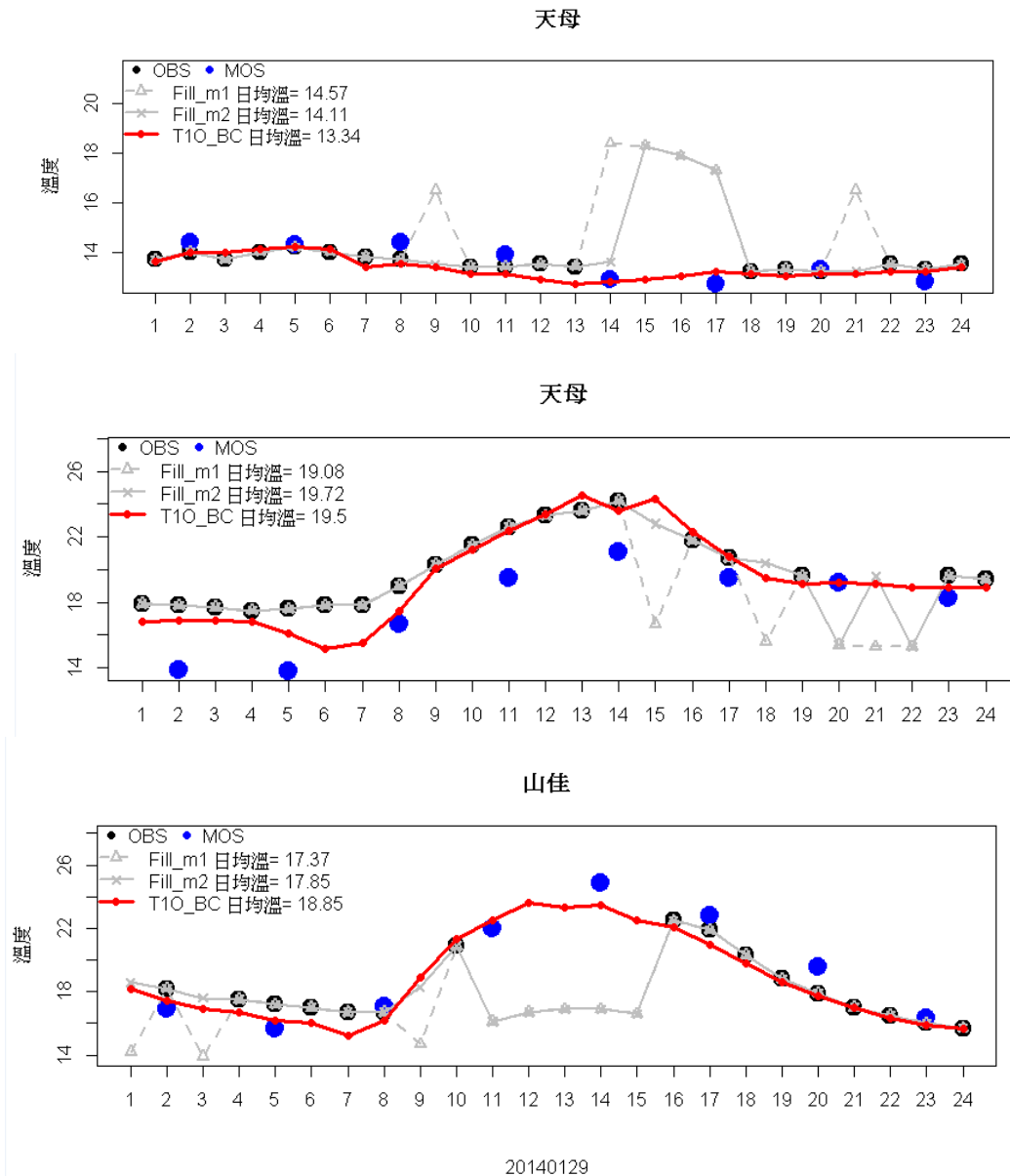


- **MOS**補遺值可協助對於 **UK** 補遺值的比對，提供一獨立結果來協助對本計畫最後挑選採用方法的可參考性評估。

利用NWP資料進行統計降尺度的Fill\_MOS法雖然也能提供修補資訊，不過此方式需要有可靠的逐小時NWP輸出值，目前並無法完全滿足對此計畫十多年長期資料補遺的需求。本年工作將以2013年及2014為例，提出示範做法，以供將來合宜時機的使用參考。

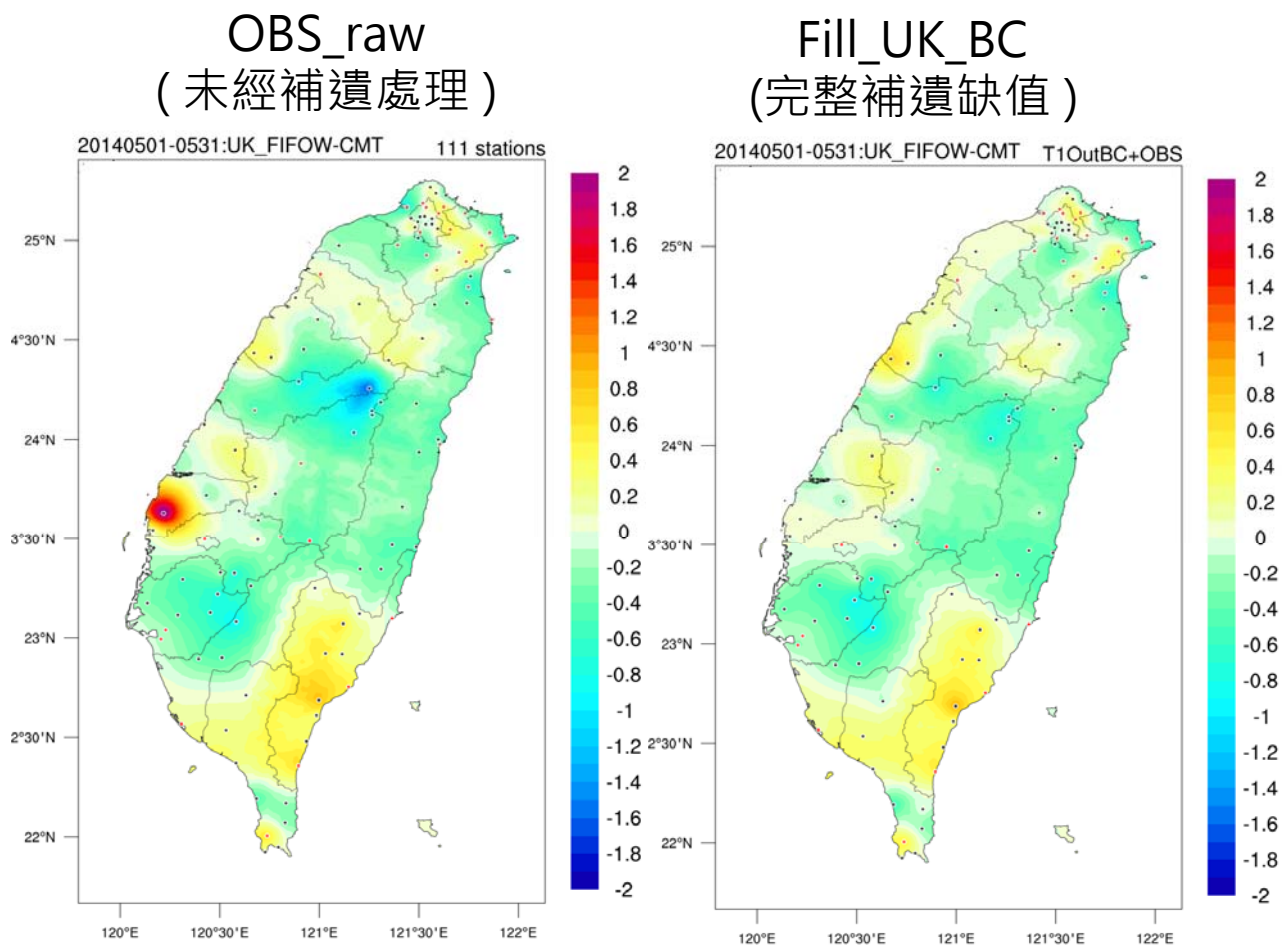
# 簡單補遺法(cmt+ano) VS 空間補遺法 (UK : T1O\_BC)

## 範例 (2)





以今年部分測站具高缺失率的5月份為例，重製經資料補遺後的網格化監測圖。

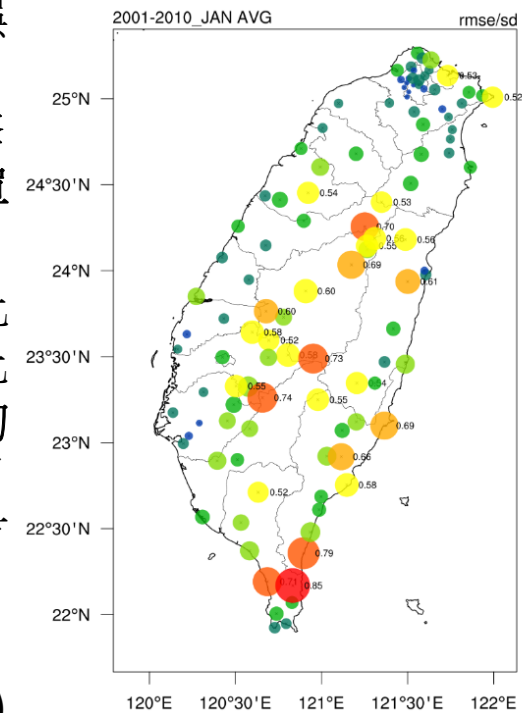


2014年5月氣溫距平圖

## 小結

- 自動站因遠地儀器維修不易及電訊傳遞有時不良等資料接收問題，其資料品質相對於人工測站較不穩定，特別在部分測站缺遺程度較高的情況下，非完整的觀測資訊將可能對網格化內插過程引入錯誤的空間統計關係。
- 為了滿足後續分析能有統計樣本均一的基礎資料，避免各站因缺遺時間的分布不同，而影響對局部特徵的代表性，我們仔細探討各種補遺方法，務使能儘可能的在一定可容許的誤差範圍內，達成使資料量均一的目標。
- 經分析，我們認為**Fill\_UK**法再經過系統性的均誤修正後的修補值合理，可大幅改進對長時間連續缺值的問題，彌補簡單補遺法(**cmt+ano**)所不能處理的問題。
- 本研究在時空的補遺皆充分依賴克利金方法，對於該網格化（或內插）方法的不確定性或可信度評估，我們也給予量化分析。因為只有在我們選用的客觀分析方法能表現出效益的前提下，才能期待其應用於補遺或網格化的適當性。經逐月分析，各月皆顯示克利金方法所得估計值能與對應觀測值有高度相關，以1月份為例，相關值在絕大多數的測站可達**0.9**以上，其中又有高比例在**0.95**以上，表示此方法確實能掌握所欲推估測站氣溫的變化趨勢。另外，從均方根誤差(**RMSE**)來量化誤差程度，各月皆顯示估計誤差皆明顯小於可代表自然變異的標準偏差(**YSD**)，表示誤差幅度確實在可接受範圍。

UK\_BC 的遮蔽測站估計誤差分析 - RMSE/YSD



# 網格化推估值受參考測站資料品質及測站數量多寡影響

選用111個參考站



去掉1個淡水站

